ENSEMBLE FORECASTING AT NMC: THE GENERATION OF PERTURBATIONS

Zoltan Toth[1] and Eugenia Kalnay

National Meteorological Center
Washington DC 20233

March 1993

NMC Office Note 391

Submitted to BAMS

[1] General Sciences Corporation, Laurel, MD

ABSTRACT

On December 7 1992, NMC started operational ensemble forecasting (Tracton et al., 1993). The ensemble forecast configuration implemented provides 14 independent forecasts every day verifying on days one through ten. In this paper we briefly review existing methods for creating perturbations for ensemble forecasting. We point out that a regular analysis cycle is a "breeding ground" for fast growing modes. Based on this observation, we devise a simple and inexpensive method to generate growing modes of the atmosphere.

The new method, "Breeding of Growing Modes", or BGM, consists of one additional, perturbed short-range forecast, introduced on top of the regular analysis in an analysis cycle. The difference between the control and perturbed six hour ("first guess") forecast is scaled back to the size of the initial perturbation and then reintroduced onto the new atmospheric analysis. Thus, the perturbation evolves along with the time dependent analysis fields, ensuring that after a few days of cycling, the perturbation field consists of a superposition of fast growing modes corresponding to the contemporaneous atmosphere, similar to the errors "bred" by the analysis cycle.

We argue that the results of the BGM method are comparable to those that can be obtained using the linear adjoint method suggested by Lorenz (1965) to generate the fastest growing perturbations, only as long as the perturbations are linear. When the perturbations are nonlinear, however, the BGM method provides a more faithful reproduction of the error growth that takes place in the analysis cycle than the Lorenz method. In particular, fast growing but low-energy and therefore mostly irrelevant perturbations, such as those associated with convection, are filtered out in both the analysis cycle and the BGM method. However, they would dominate the spectrum of fast growing perturbations in the Lorenz method if the physical parameterizations of the primitive equations model were faithfully included. This, and its low cost and simplicity, are advantages that the BGM method has over Lorenz' method, which is also used experimentally at other centers (ECMWF, NCAR). Preliminary results indicate that ensembles of just two BGM forecasts achieve better results than much larger Monte Carlo or Lagged Average Forecast ensembles. Therefore, the operational ensemble configuration at NMC is based on the BGM method to generate efficient initial perturbations.

1. Why operational ensemble forecasting?

On 7 December 1992, the National Meteorological Center replaced the single 10-day global medium range forecast (MRF) which was run daily at 00Z, by an ensemble of four 12-day forecasts, plus an extension to 12 days of the Aviation 3-day forecast run at 12Z (Tracton et al, 1993). The operational configuration implemented at that time is such that there are 14 forecasts, originating from analyses within the most recent 48 hours, that verify over the same 10 day period. It replaces the previous configuration, where only one operational forecast and one experimental forecast were available for the 6-10 day forecast range. In order not to increase the total use of the CRAY YMP supercomputer, which is already saturated, a compromise had to be found, where the resolution of the MRF was reduced beyond day 6 from triangular truncation T126 (equivalent to a gaussian grid resolution of 105 Km), to T62 (equivalent to 210 Km). It was found, however, that the reduction of resolution did not affect significantly the quality of the forecasts as long as it was performed after the first 5 days of the forecast (Tracton et al, 1993).

The replacement of single operational forecasts by an ensemble of operational forecasts reflects explicitly the recognition that the atmosphere is a chaotic system. As pointed out by Lorenz (1962), even an infinitesimally small perturbation (as would be produced, for example, by the "wings of a butterfly") introduced into the state of the atmosphere at a given time will result in an increasingly large change in the evolution of the atmosphere with time, so that after about two or three weeks, the trajectories of the perturbed and the original atmosphere would be completely different.

Lorenz' discovery led to the emergence of a new discipline, dynamical systems theory[1], and to the realization that many apparently deterministic systems, like the atmosphere and its numerical models, are ***also*** chaotic: arbitrarily small initial perturbations evolve into large differences with time. As far as real physical systems are concerned, their state can never be measured exactly: for example, we know that our analyses of the atmosphere contain errors whose magnitude can only be estimated. The analysis errors are due to errors in measurements and in the first guess, lack of complete data coverage and approximations in our analysis techniques. Even with a perfect model of the atmosphere, the skill of our forecasts would degrade to zero within a few weeks. However, if we are willing to run an ensemble of forecasts from slightly perturbed initial conditions, then averaging the ensemble should filter out some of the unpredictable components of the forecast, and the spread among the forecasts should provide some guidance on the reliability of the average skill.

The main subject of this paper is one of the central questions of ensemble

---

[1] Some of the mathematical formalism had been already developed by Poincare and his contemporaries in the late 19th century.

forecasting, namely: ***How can we generate realistic perturbations for the initial conditions?*** If we want for some of our perturbed forecasts to follow the true evolution of the atmosphere closer than the control forecast, we need perturbations that grow as fast as the errors in our model forecasts. In other words, if we neglect model deficiencies, the true atmospheric evolution should be a plausible member of a realistic ensemble.

We will argue in Section 3 that the analysis cycle contains well organized, fast growing errors associated with the instabilities of the atmospheric state, and that effective perturbations should, at least locally, share the shapes of these errors. Since the number of degrees of freedom governing the evolution of the large-scale atmospheric flow is large, it is very unlikely that random perturbations would project substantially on such fast growing modes. Without a systematic selection of fast growing perturbations, ensemble forecasting will be much less effective.

In section 2 we give a brief overview of the methods that have been used so far to create ensemble perturbations. In Section 3 we discuss how fast growing errors are introduced and maintained in a conventional analysis cycle. In Section 4 we briefly discuss a method in which fast growing perturbations are obtained in a "breeding cycle" of growing modes (Toth and Kalnay, 1991a). Section 5 gives an overview of the results of ensemble forecasts based on the bred growing modes (BGM) method. A discussion of the results and of other applications for which the breeding method can also be used is given in Section 6. A companion paper (Tracton et al., 1993) includes further details of the new NMC ensemble operational configuration, including a discussion of applications.

2. Generation of ensemble perturbations: An overview

The idea of ensemble forecasting was first introduced by Leith (1972). He suggested a "Monte Carlo" forecasting method, where the initial conditions of a control forecast are perturbed by adding small random numbers. These perturbed analyses are the initial conditions for an ensemble of forecasts. In recent years it has become apparent that simple Monte Carlo Forecasting (MCF) is not the best way of making ensemble forecasts. This is because random perturbations require a few hours or even days before they organize into dynamically unstable modes that grow as fast as forecast errors are observed to grow.

Another method, Lagged Average Forecasting (LAF) was suggested by Hoffman and Kalnay (1983). This scheme, which like MCF has also been widely used, takes advantage of operational forecasts launched before "today" as members of an ensemble. LAF perturbations are basically "realistic" short term forecast errors, i.e. difference fields between a forecast and an analysis (see schematic Fig. 1). Let us assume that the analysis at an earlier time contains both the growing and non-growing type of errors when compared to "truth". By the end of a short-range forecast, the proportion of growing errors will, by definition, be enhanced in the difference between the forecast and our new estimate of truth, the subsequent analysis. This is the reason why LAF perturbations grow faster than random perturbations of the same magnitude. However, LAF ensemble forecasting has the disadvantage that earlier or "older" forecasts have much larger "perturbations", and hence are considerably less skillful than later or "younger" forecasts. This problem can be partly alleviated by either using different weights for different members of the ensemble (Hoffman and Kalnay, 1983), or by scaling back the larger errors to a reasonable size (SLAF[1], Ebisuzaki and Kalnay, 1991).

To further increase the growing component in the perturbations, Kalnay and Toth (1991a) used the difference between short-range forecasts (SRFD), started at earlier times but verifying at the initial time of the ensemble (Fig. 1). Here, growing errors further dominate the difference between the forecasts, and, unlike LAF or SLAF, no new random errors are introduced by the latest analysis. Experiments performed by Toth and Kalnay (1991b) showed that there was a clear increase in the growth rate of perturbations from MCF to SLAF and from SLAF to SRFD, which was accompanied by an increase in the quality of the ensemble, measured by the skill of the mean of the ensemble forecasts. The faster the initial error growth, the better the perturbation. This is because if there are fast growing errors in the analysis (and we argue in the next section that there are) then we have no chance of capturing a trajectory close to the true evolution of the atmosphere, unless we perturb the analysis along those growing errors.

Theoretically, Lorenz (1965) showed that in a linear sense, the fastest growing perturbations for a given period of model integration can be obtained as the eigenmodes of the product A^*A of the linear model propagator $A(t_1,t_0)$ between a time t_0 and t_1, with its adjoint A^*. For small models, the eigenvectors of A^*A, (singular vectors of A), can be directly obtained, and ***constitute optimal perturbations for ensemble forecasting, for this particular period (t_0,t_1) and as far as the linear assumption holds.*** A number of researchers (e.g., Lacarra and Talagrand, 1988, Farrell, 1989, Molteni and Palmer, 1992, Mureau et al, 1992, Tribbia, pers. comm., Borges and Hartmann, 1992) have applied Lorenz' method to ensemble forecasting as well as to studies of atmospheric instability.

---

[1] In scaled Lagged Average Forecasting (SLAF), proposed recently by Ebisuzaki and Kalnay (1991), the LAF perturbations are divided by an "age" factor, resulting into similarly sized realistic perturbations. For example, a 12-hour forecast error perturbation could be divided by two and used as if it was a 6-hour forecast error. They also suggested that the perturbations could be both added and subtracted from the control (i. e., latest) analysis. In this way, it is possible to create a 17-member ensemble from just the latest two days of 6-hour analysis cycle. Unlike MCF or LAF, this scheme resulted in average forecasts that verified as well or better than the control forecast (the forecast launched from the latest analysis) even after only 12 hours into the forecast.

For large atmospheric models such direct determination of the fastest growing modes is computationally very expensive. For this reason Molteni and Palmer (1992), and Palmer et al. (1992), have chosen to use Lorenz (1965) method to find the eigenmodes of A^*A, not for the T63/19-level ECMWF model they were using for extended range forecasting, but for a T21/3-level quasi-geostrophic model, with much fewer degrees of freedom and simpler physical processes. More recently, experiments were also done at ECMWF obtaining the fast growing modes from the T63/L19 primitive equations model, but the lack of physical parameterizations in the linear model and its adjoint introduced further difficulties (Buizza, 1992).

At NMC we have tried to find an efficient, inexpensive method to create realistic perturbations which could represent well the errors which are actually present in the analysis cycle. The method we have developed, denoted "breeding of growing modes", or BGM, (Toth and Kalnay, 1991a), has been applied in the operational implementation of 7 December 1992 (Tracton et al, 1993). As we discuss in the next section, the analysis cycle is a "breeding ground" for fast growing errors, and the new method presented in Section 4 is designed to mimic it.

3. Errors in the analysis cycle

As indicated in Section 2, it has become apparent in recent years that simple Monte Carlo Forecasting is not an effective way to do ensemble forecasting. To understand why this is the case, we should consider the analysis cycle, which is the basis for modern operational numerical weather prediction.

In a typical 6-hour operational analysis cycle, a global model starts from initial conditions given by a previously completed atmospheric analysis, and is integrated for a short (6 hour) forecast. The 6-hour forecast serves as a ***first guess*** for the next analysis, which is the statistical combination of the first guess with observations collected in a +/- 3-hour window centered at the time of validity of the forecast. This cycle is run 4 times a day, every day (Fig. 2). When analysis schemes were introduced into meteorology, climatology or persistence were also used as a first guess in the analysis (e.g., Gandin, 1963), rather than a short range forecast. The use of a model forecast as a first guess in the analysis cycle (also denoted 4-dimensional data assimilation because of the time dimension introduced by the model) has resulted in major improvements in operational forecast skill in the last 10 to 15 years (e.g., Kalnay et al, 1990). The analyses (initial conditions) have much smaller overall errors and are far superior to those that could be obtained by using persistence or climatology as a first guess. This is because the model acts as a transporter of information from data rich to data poor regions, and therefore provides a good estimate of the actual state of the atmosphere even for those regions or parameters for which there are no observations (e.g., Charney et al, 1969).

The total error in the analysis has therefore been decreased by the introduction of a model first guess. However, the ratio of fast growing errors to the total error must have increased. We know that the analysis is only a close approximation of the true state of the atmosphere. The errors in the analysis (analysis minus truth) are not known, but they must contain both fast-growing, high-energy perturbations, such as baroclinically unstable modes, and slow- or non-growing, low-energy perturbations, such as gravity waves. The random, non-growing part of the error comes mainly from observational errors. Regarding the growing errors, they originate mainly from the forecasts. When a 6-hour forecast is run from an analysis, which has both types of errors, the fast-growing, high-energy perturbations will grow faster, attain relatively large amplitudes, and dominate the error at the end of the forecast.

The use of the new data collected in the +/- 3-hour window will only reduce the size of the error (Fig.2), but in general will not completely remove the fast-growing errors, which remain present and evolve and amplify again in consecutive analyses. Hence, the high ratio of fast growing errors in the analysis is introduced and maintained through the successive use of short-range forecasts in the analysis cycle. For this reason, there is a "natural selection of fast growing errors" in the analysis, and the error growth in the short range forecasts is generally high.

Ensemble forecasting can be considered as an attempt to simulate this error growth, i.e., the process by which the trajectory of the atmosphere in the forecast moves away from the trajectory of the observed atmosphere. If we try to do ensemble forecasting with ensemble members which differ from each other by random perturbations in the initial conditions, we stand no chance of simulating the true departure between the model and the atmospheric trajectories. The fast growing modes present in the analysis have well organized structures which are highly dependent on the atmospheric flow, i.e., they are instabilities corresponding to the "flow of the day". Random perturbations will in general project mostly on slowly growing and even decaying modes for the first few hours or days. Eventually, after a day or two, fast growing modes will become organized, and dominate any further error growth, but by then the model forecast has already moved away from the atmosphere much faster, and in other phase directions, than two forecasts with randomly perturbed initial conditions. For this reason we have developed a method (Toth and Kalnay, 1991a) denoted "breeding" of fast growing perturbations, in which we try to mimic the natural selection of fast growing errors that takes place in the analysis cycle.

4. The breeding of fast-growing perturbations

This simple and inexpensive method (Toth and Kalnay, 1991a) consists of the following steps: a) add a small perturbation to the atmospheric analysis, b) integrate the model for 6 hours from both the unperturbed (control) and the perturbed initial condition, c) subtract the 6-hour control (analysis cycle) forecast from the perturbed forecast, and d) scale down the difference field so that it has the same size (in an rms sense) as the initial perturbation. This perturbation is now added to the following 6-hour analysis, as in a), and the process is repeated forward in time (Fig. 3). By construction, this method selects ("breeds") the modes that grow fastest during the cycle, in a way similar to that which occurs in the analysis cycle itself.

We have performed experiments with the simple 3-variable Lorenz (1963) model confirming that the breeding method results in growing perturbations very similar to those obtained as eigenmodes of A^*A (singular modes of A) whenever perturbations are growing fast. The growth rates obtained by both methods are almost identical. This is because the propagator $A(t_1,t_0)$ is in effect applied ***once*** during the interval t_0 to t_1, and, by definition, it is the singular modes (not the eigenmodes) of A that grow fastest during a ***single*** application of A.

The use of the breeding cycle has an additional important advantage over the Lorenz (1965) method in simulating the breeding of fast growing errors that takes place in the analysis cycle. Both the analysis cycle and the breeding cycle use ***nonlinear*** models with a complete set of physical parameterizations, whereas the adjoint method of Lorenz uses a ***linear*** version of the model and its adjoint, in practice with only minimal physical parameterizations of subgrid-scale processes. Therefore, the Lorenz method does not include the non-linear effects that lead to saturation of the "errors" or the perturbations. For example, modes associated with unstable convection have an exceedingly fast growth rate (doubling time of less than one hour). However, these modes saturate at an energy level much smaller than the baroclinically unstable modes, which dominate the growing errors in the analysis (Fig. 4). Because of linearity, the Lorenz (1965) method may select many fast growing but low-energy modes, which are in practice irrelevant for the analysis problem[2]. Because of saturation, the non-linear breeding method, on the other hand, will filter out these irrelevant modes from the initial perturbations in the same way that the analysis cycle does. Finally, since the modes obtained by breeding are already balanced, there is no need to introduce non-linear normal mode initialization, which has a strong effect in the linear adjoint algorithm (Buizza, 1992).

A "breeding cycle" of growing modes was implemented at NMC on the quasi-operational 18-level T62 analysis cycle in May 1992. The perturbation amplitude used is about 10% of the natural variability of the atmosphere in an rms sense, which is also

---

[2] This problem should not be very apparent in a quasi-geostrophic model, but should appear in a primitive equation model, where the physics, or lack of appropriate physics in the linear and adjoint models, will lead to fast growing singular vectors that need to be omitted from the ensemble of perturbations. Such domination of growing convective modes was observed in the 6-hour breeding cycle when the size of the perturbation field was chosen to be sufficiently small (0.1% of the atmospheric variability), so that those modes grew within their linear range.

the estimated size of the analysis errors. Since the period of integration used is 6 hours, the same as the analysis cycle, the unperturbed forecast is already available, so that the net cost of creating these fast growing modes is only one day of model integration per day. Experiments have shown that the results are not very sensitive to the length of the integration, at least for periods between 6 hours and two or three days. (However, as indicated in the footnote of the previous paragraph, and suggested by the schematic Fig. 4, for much smaller perturbation amplitudes, convective instabilities rather than slower but energetic baroclinic instabilities become the fastest growing modes dominating the breeding cycle).

Experimental results so far indicate that there is a qualitative relationship between the growing modes of a given day and the actual observed forecast errors for the same day. However, the breeding cycle may be capturing only ***a subset of the relevant fastest gowing modes***. Experiments show that it is possible to capture additional growing modes by using ***multiple breeding cycles started from different initial perturbations***. In this case, after a few days of breeding, we find that two independent cycles share many (over half) of the local modes, but that there are regions with different modes, presumably because there are competing growing modes with ***similar growth rates but different initial projections in the two cycles.***

It should be noted that it is possible to generate additional breeding cycles for ensemble forecasting performed daily at no additional cost. This can be done by defining the growing modes as the difference between a pair of forecasts at day 1 that were obtained by adding and subtracting the previous day's growing mode to the analysis at that time. ***Pairs of extended-range forecasts, started with independent perturbations, can therefore maintain their own breeding cycles*** (Fig. 5a). In this configuration, efficient perturbations are generated without any extra computational cost (beyond running the extended-range forecasts). If single additional perturbed members of the ensemble are desired, rather than twin pairs, ***each forecast can also maintain its own breeding cycle with respect to the control forecast*** (Fig. 5b).

5. Ensemble forecasts using the bred growing modes (BGM)

As further discussed in Tracton et al. (1993), ensemble forecasting has three major objectives: a) improve the skill, by reducing the non-linear error growth and averaging out unpredictable components, b) predict the skill, by relating it to the agreement among ensemble forecast members, and c) provide an objective basis for casting forecasts in a probabilistic form. The latter two objectives, which involve higher moments, are hard to verify without large numbers of forecasts. However, the impact of ensemble prediction on forecast skill can be easily verified, and can be used to discriminate among different methods for generating perturbations.

Since we believe that the growing modes obtained by the breeding method are a very good representation of the "errors of the day" present in the analysis cycle, we have followed a minimalistic ensemble forecast approach: we generated two-member ensembles using the bred modes as perturbations. The ensemble contains two initial perturbations, obtained by adding and subtracting to the analysis our growing mode estimate for the same day.

Five day forecasts were run with the perturbed analyses, and the skill of the mean of the two perturbed forecasts was compared with the skill of the control forecast and also with the skill of "benchmark" ensembles of much larger size generated by other methods (MCF and SLAF). The size of the initial perturbation was kept constant throughout these experiments at a level of 10% of the climatological variance. All the experiments described below were performed with the T62/18 levels version of the NMC global model (Kanamitsu et al., 1991).

The results of experiments performed over several months show, first, that ***the five-day forecast skill of the mean of the twin ensemble based on the bred growing modes is higher than the skill of the control in about 80% of the forecasts for both the Northern and the Southern Hemispheres***. Experiments completed for February and March 1992, show that the average improvements in the anomaly correlation were 2% and 3% for the Northern and Southern Hemispheres, respectively.

These early results have been confirmed by those obtained since the operational implementation of ensemble forecasting on 7 December 1992. Table 1 shows that the forecasts started at high resolution (T126) are considerably more accurate than those started at lower resolution (T62), but that the mean of the twin BGM lower resolution forecasts has skill comparable to that of the high resolution control forecasts. It is important to note that running twin BGM forecasts requires twice as much computational resources as running a single lower resolution forecast, while doubling the resolution requires 8-10 times more resources. In this sense, running twin forecasts is a ***much more cost effective way of enhancing the skill than increasing the horizontal resolution.*** Moreover, the BGM forecast has substantially more skill than the 3-member LAF forecasts, which include "older" forecasts with higher error levels[3].

Second, ***the skill of the twin esemble is comparable, or even better than the skill of much larger ensembles using previous methodologies to obtain perturbations***. In Table 2, we show the anomaly correlation for a set of control 5-day forecasts and for the mean of the different ensemble forecasts. Both the SLAF and MCF ensembles contain 8 pairs, each of which is generated by adding and subtracting a different perturbation, whereas a single pair is used with the bred growing modes (BGM). In SLAF, the perturbations are short-range forecast errors (Ebisuzaki and Kalnay, 1991),

---

[3]Note that the scaled LAF method (SLAF) would not suffer from this handicap, but would require additional model integrations.

while in MCF, the perturbations are combinations of randomly chosen earlier analysis fields (Kalnay and Toth, 1991b). The scores, computed and averaged for 6 independent cases (3 individual days for both the Northern and Southern extratropics), show that ***with only two members the BGM forecasts achieve the same improvement over the control as the SLAF method with 16 forecasts, and in turn, that the SLAF is considerably better than what could be achieved with MCF perturbations***. This latter result is not unexpected, since the SLAF perturbations, like LAF perturbations, are based on short term forecast errors, and, as discussed in Section 2, are more likely to contain realistic growing "errors of the day" than random Monte Carlo perturbations.

In order to have a clearer estimate of the difference between a single pair of BGM and MCF ensemble, we performed another set of forecasts for 26 February 1992. We ran two independent breeding cycles, started weeks earlier from different random initial perturbations. As indicated before, such independent breeding cycles after a few days end up sharing about half of the growing modes, and have different modes in the rest of the hemispheres. We ran 5-day forecasts using both breeding cycles, and averaged the score of each twin pair. For comparison, we ran 18 twin pairs of Monte Carlo perturbations, and averaged the score of each individual pair. We also computed the score for the mean of all 36 Monte Carlo forecasts.

As Table 3 shows, ***a single bred growing mode twin forecast achieves as much improvement over the control forecast as the eighteen times larger ensemble of 36 MC forecasts***. Moreover, the average improvement from single MC twin forecasts is only about one fourth of that obtained with BGM twins.

To illustrate the typical shapes of a field of growing modes in the BGM scheme, and the result of twin ensemble forecasting, we show in Figs. 6-8 an example of a successful BGM forecast. The initial conditions for this case, 00Z 15 February 1992, and the bred growing modes obtained for the same day are shown in Figs. 6a and 6b, respectively. The results are presented in terms of the streamfunction at mid-level of the atmosphere, which is similar to the 500 hPa geopotential field in the extratropics, but resolves better the tropical circulation.

Note that the growing modes are associated with specific features of the atmospheric circulation on that day. For example a dipole pattern of modes near the Caspian Sea in Fig. 6b is associated with a traveling low in Fig. 6a in the same area, or the train of short waves off the West Coast of North America are accompanied by a cut-off low.

Fig. 7a shows the verification analysis corresponding to 00Z 20 February 1992, Fig. 7b the 5-day control (single) forecast, and Fig. 7c the mean of the BGM twin ensemble. Several streamlines have been highlighted in areas where the control and the BGM twin average forecasts differ substantially in both the Northern and the Southern hemispheres. In most areas of disagreement (but not in all) the twin BGM average forecast is closer to the analysis than the control forecast. This subjective statement is substantiated by the objective scores (anomaly correlation), indicated in Table 4. It should be noticed that ***the improved forecast verification is not a result of general smoothing of small scale features***, as is the case when large number of ensemble members are averaged, or when a time average is performed.

Finally, we note that when we verified over a large number of cases each individual member of the BGM ensembles against the analysis, we found that ***in 67% of the forecasts, at least one member of the twin growing mode ensemble pair was better than the control forecast, whereas this ratio was only 17% for random Monte Carlo perturbations over the Northern Hemisphere***. To further illustrate this point, we show in Fig. 8 the negatively and positively perturbed 5-day forecasts for the above case of 20 February 1992. The corresponding skill scores are listed in Table 4. The high ratio of ensemble members that are superior to the control forecast is a clear indication that our growing mode estimates do project onto the actual error field of the control analysis and thus provide excellent initial perturbations. This also suggests that it should be possible to take advantage of the BGM cycle to improve the analysis.

6. Discussion

We have presented a simple method, denoted Breeding of Growing Modes (BGM) to create fast growing perturbations that can be added to the analysis as initial perturbations for ensemble forecasting. These perturbations have been used in the ensemble forecasting scheme that became operational at NMC on 7 December 1992. The details of the ensemble configuration, along with possible applications, are presented in Tracton et al (1993).

We have shown that the new method is simple, inexpensive, and, when compared with previous methods for ensemble forecasting, like Monte Carlo and Lagged Average Forecasting, provides more effective forecasts. We have argued that these advantages come from the fact that the breeding method mimics the operational analysis cycle, which is itself a breeding ground for fast growing instabilities highly dependent on the details of the circulation at the initial time. For this reason a random MCF scheme cannot compete with the BGM since random perturbations will not in general grow at the initial time like the initial errors in the forecast grow.

We argued that the results of the BGM method are comparable to those that can be obtained using the adjoint method suggested by Lorenz (1965) to generate the fastest growing perturbations, only as long as the perturbations are linear. When the perturbations are nonlinear, however, the BGM method provides a more faithful reproduction of the error growth that takes place in the analysis cycle than the Lorenz method. In particular, fast growing but low-energy and therefore mostly irrelevant perturbations, such as those associated with cumulus convection, are filtered out in both the analysis cycle and the BGM method, whereas they would dominate the spectrum of fast growing perturbations in the Lorenz method if the physical parameterizations of the primitive equations model were faithfully included.

It is clear from the above that effective perturbations for ensemble forecasting crucially depend on the instabilities of the flow. This is because in the analysis cycle, the very same instabilities, through the repetitive use of short-range forecasts that enhance the growing errors, establish in turn error patterns in the control analysis that are "ready to grow". Research related to ensemble forecasting and atmospheric analysis are hence closely related and may mutually benefit from each other, as suggested in the last paragraph in section 5.

Preliminary analysis experiments with Lorenz' 3-variable model and also with the full NMC global system indicate at least two possible ways to use the bred modes to improve the analysis. First, the growing modes can be used to estimate the forecast error covariance, a task which is normally very difficult because it depends on the availability of sparse rawinsonde observations, burdened with random errors, to estimate very short-range forecast errors. Since the growing modes are complete 3-dimensional fields that describe the shape of the fastest growing errors that appear in the short range forecasts, their outer product can be computed without difficulty, and should provide an efficient and accurate parametrization for improving the forecast error covariance used at the next data time (D. Parrish, pers. comm.). Second, it may be possible to improve the analysis by locally modifying the first guess towards the observations along the direction of the growing modes.

In addition to analysis and ensemble forecasting, the BGM method should also be useful for other applications such as forecast of the skill and stability studies of the observed or modeled atmosphere. The most unstable, fastest growing modes can be estimated by breeding, either for a sequence of analyses, representing the evolution of the atmosphere, or for a GCM long run. The growth rate, statistical and dynamical properties and structure of the most unstable perturbations can be derived from these fields. For example, an examination of the cross-section of the zonally averaged rotational kinetic energy of the growing modes clearly shows that they are able to penetrate the winter stratosphere, but not the summer stratosphere, as might be expected from Charney-Drazin theory.

We plan to present a detailed description of the BGM methodology, characteristics of the growing modes, and their application to the analysis and to ensemble forecasting in forthcoming papers.

7. Acknowledgements

We are most grateful to M. Kanamitsu, H. Juang, M. Iredell and S. Saha for their essential help with both the research and computational aspects of our work, and with its operational implementation. We also had many stimulating discussions with D. Parrish, J. Derber, J. Purser, H. van den Dool, W. Ebisuzaki, T. Palmer, J. Ahlquist, and J. Anderson. J. Irwin and other members of the Automation Division were extremely helpful in the operational implementation of the breeding cycle. S. Tracton provided the data for Table 1. The encouragement and support of R. McPherson are much appreciated.

8. References

Borges, M. D., and D. L. Hartmann, 1992: Barotropic instability and optimal perturbations of observed non-zonal flows. *J. Atmos. Sci.*, **49**, 335-354.

Buizza, R., 1992: Unstable perturbations computed using the adjoint technique. Research Department Technical Memorandum No. 189, ECMWF. Available from: ECMWF, Reading, UK.

Charney, J. G., M. Halem, and R. Jastrow, 1969: Use of incomplete historic data to infer the present state of the atmosphere. *J. Atmos. Sci.* **26**, 1160-1163.

Ebisuzaki, W., and E. Kalnay, 1991: Ensemble experiments with a new Lagged Analysis Forecasting scheme. Research Activities in Atmospheric and Oceanic Modelling, Report No. 15, WMO.

Farrell, B. F., 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.* **46**, 1193-1206.

Gandin, L. S., 1963: Objective analysis of meteorological fields. Leningrad, Gidrometeoizdat, in Russian. (English translation, Israel Program for Scientific Translations, Jerusalem, 1965, 242 pp.)

Kalnay, E., M. Kanamitsu and W. E. Baker, 1990: Global Numerical Weather Prediction at the National Meteorological Center. *Bull. Am. Meteor. Soc.* , **71**, 1410-1428.

Kalnay, E., and Z. Toth, 1991a: Estimation of the growing modes from short range forecast errors. Research Highlights of the NMC Development Division: 1990-1991, p. 160-165. Available from NMC, Washington DC 20233.

Kalnay, E., and Z. Toth, 1991b: Efficient selection of Monte Carlo forecast ensemble members. Research Highlights of the NMC Development Division: 1990-1991, p. 157-159. Available from NMC, Washington DC 20233.

Kanamitsu, M., J. C. Alpert, K. A. Campana, P. M. Caplan, D. G. Deaven, M. Iredell, B. Katz, H.-L. Pan, J. Sela, and G. H. White, 1991: Recent changes into the Global Forecast System at NMC. *Weather and Forecasting*, **6**, 425-435.

Lacarra J. F., and O. Talagrand, 1988: Short range evolution of small perturbations in a barotropic model. *Tellus*, **40A**, 81-95.

Lorenz, E. N., 1963: Deterministic non-periodic flow. *J. Atmos. Sci.*, **20**, 130-141.

Lorenz, E. N., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321-333.

Molteni, F. and T. N. Palmer, 1992: Predictability and finite-time instability of the Northern Winter circulation. *Quart. J. Roy. Met. Soc*, **119**, 269-298.

Mureau, R., F. Molteni, and T. N. Palmer, 1992: Ensemble prediction using dynamically conditioned perturbations. *Quart. J. Roy. Met. Soc*, **119**, 299-323.

Palmer, T., F. Molteni, R. Mureau, R. Buizza, P. Chapelet and J. Tribbia, 1992: Ensemble Prediction. ECMWF Research Department Tech. Memo. No 188. Available from ECMWF, Reading, UK.

Toth, Z., and E. Kalnay, 1991a: Estimating the growing modes of the atmosphere: The breeding method. Research Highlights of the NMC Development Division: 1990-1991, p. 439-443. Available from NMC, Washington DC 20233.

Toth, Z., and E. Kalnay, 1991b: Ensemble forecasting: Are two members enough? Research Highlights of the NMC Development Division: 1990-1991, p. 435-438. Available from NMC, Washington DC 20233.

Tracton, S., H. van den Dool, E. Kalnay, Z. Toth, M. Iredell and J. Irwin, 1993: Ensemble forecasting at NMC: Operational Implementation. To be submitted to Weather and Forecasting.

Table 1: Comparison of the average 6-10 day 500 hPa forecast skill (Northern Hemisphere extratropics) obtained using subsets of the operational ensemble forecasts. The verification period is 13 December 1992 through 13 January 1993. AC represents percentage anomaly correlation, and RMS the rms error scaled by the climatological variability. HR represents forecasts performed at 00Z, using the high resolution T126 model and analysis for the first 6 days, and then extended to 12 days with T62 resolution. LR represents forecasts performed at 00Z using the lower resolution T62 analysis and the T62 model throughout the 12 days of integration. CTL represents control forecasts started from unperturbed 00Z analysis. BGM represents the mean of the twin forecasts started from control analyses perturbed with positive and negative bred growing modes from the breeding cycle. LAF are 3-member lagged averaged forecasts obtained from the latest 3 00Z analyses, using uniform weights. The last entry is the average of all the 4 forecasts performed using the latest 00Z analyses.

FORECAST TYPE	AC	RMS
CTL(HR)	61.3	0.88
CTL(LR)	56.3	0.96
BGM(LR)	60.2	0.88
LAF(HR)	59.6	0.85
LAF(LR)	53.5	0.94
CTL(HR)+CTL(LR)+2BGM(LR)	62.6	0.84

Table 2: Verification of the mean of different ensemble forecasts generated by various methods. The anomaly correlation values of 5-day forecasts are averaged for 8, 17 and 22 January 1992 and over the Northern and Southern Hemisphere extratropics.

ENSEMBLE TYPE	Control	BGM Twin	SLAF	MCF
NUMBER OF FORECASTS	1	2	16	16
ANOMALY CORRELATION	65.7	67.3	67.3	66.5

Table 3: Same as Table 2, except that the bred growing modes BGM Twin includes the average of the scores of two pairs of independent bred growing modes twin ensembles, while MCF Twin gives the average of the scores of 18 randomly perturbed twin forecasts. The MCF gives the score of the mean forecast obtained using all 36 members of the MC ensemble. All results are for February 26, 1992, with the NH and SH scores combined.

ENSEMBLE TYPE	Control	BGM Twin	MCF Twin	MCF
NUMBER OF FORECASTS	1	2	2	36
ANOMALY CORRELATION	72.0	74.5	72.8	74.4

Table 4: Anomaly correlation values for 5-day forecasts verified on 20 February 1992, over the Northern and Southern Hemisphere extratropics separately. Next to the score for the control forecast are the scores for the two individual forecasts perturbed by adding and subtracting the BGM modes and also for the mean of the two perturbed forecasts (BGM twin).

FORECAST		BGM neg. perturb.	Control	BGM pos. perturb.	Mean of neg.& pos.
Anomaly	NH	59	58	68	70
Correlation	SH	71	66	64	70

Figure captions:

Fig. 1: Schematic of the creation of Lagged Average Forecasting (LAF) and Short-Range Forecast Difference (SRFD) perturbations. Note that the LAF perturbation includes not only the short range forecast errors, but also the random errors of the latest analysis, whereas the SRFD perturbation is not affected by the random errors of the latest analysis. This results in a significant reduction of the random errors and therefore in a higher growth rate for the SRFD perturbations.

Fig. 2: Schematic of the 6-hour analysis cycle. On the vertical axis indicated are differences between the true state of the atmosphere (or its observational measurements, burdened with random errors) and the analysis or forecasts of it. Note that the difference between a forecast and the true state of the atmosphere (or the observations) increases with time, due to the growing type of errors in the initial analysis.

Fig. 3: Schematic of the 6 hour breeding cycle. Note that the breeding cycle depends on the analysis cycle but does not affect it. A small arbitrary perturbation is introduced on the control analysis initially. After a 6-hr integration, the difference between the control and perturbed forecasts is scaled back to the size of the initial perturbation and this difference field is now added onto the new analysis. After 3-4 days of cycling, the perturbation is dominated by growing modes due to the "natural selection" of fast growing perturbations.

Fig. 4: Schematic of the time evolution of the rms amplitude of high energy baroclinic modes, and low energy convective modes. Note that though initially growing much faster than the baroclinic modes, convective modes saturate at a substantially lower level. These modes are therefore insignificant for the analysis/ensemble perturbation problem since the errors in the control analysis (dashed line) are ***much larger*** than the convective saturation level.

Fig. 5a: Schematic showing how ensemble pairs of perturbed extended-range forecasts can maintain their own breeding cycle. An arbitrary perturbation is introduced with a negative and positive sign on the initial control analysis. The following day the difference between these forecasts at one day lead time serves as a new perturbation. After 3-4 days, the perturbations are dominated by fast growing modes, just as in the breeding cycle. Fig. 5b: Schematic showing how a single perturbed forecast can maintain its own breeding cycle by defining the perturbations with respect to the control forecast.

Fig. 6: a) Analysis corresponding to the initial conditions of 15 February 1992. b) Bred growing modes for 15 February 1992. Shown are the streamfunction values at sigma layer 9 (around the 500 hPa height level; displayed values are in 10^7 m^2s^{-1}).

Fig. 7: a) Verifying analysis for 20 February 1992. b) 5-day control forecast. c) 5-day mean forecast from the BGM twin forecasts. For further details, see Fig. 6.

Fig. 8: a) 5-day forecast from the BGM positive perturbation. b) 5-day forecast from the BGM negative perturbation. For further details, see Fig. 6.

LAF and SRFD perturbations

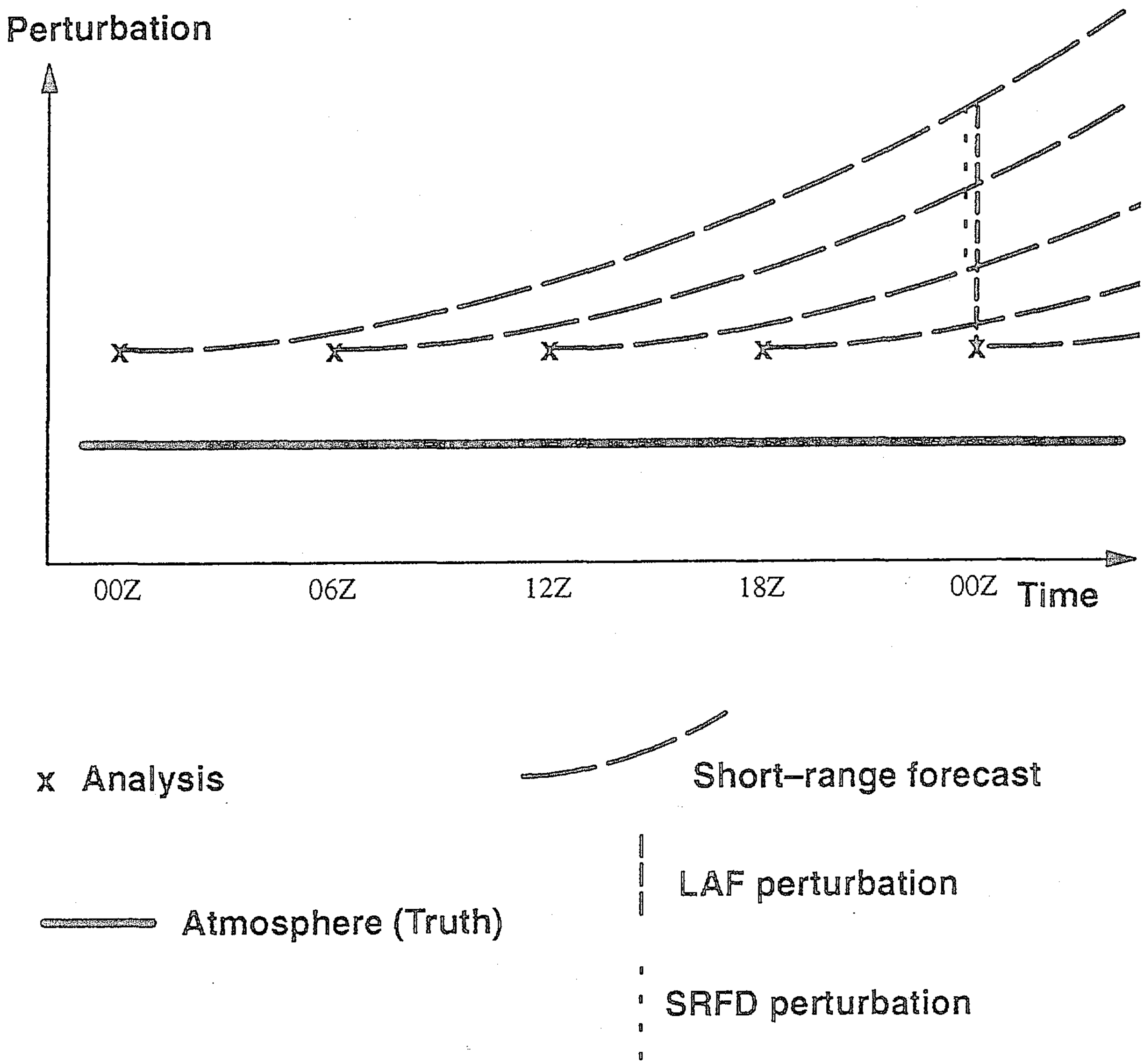


Fig. 1

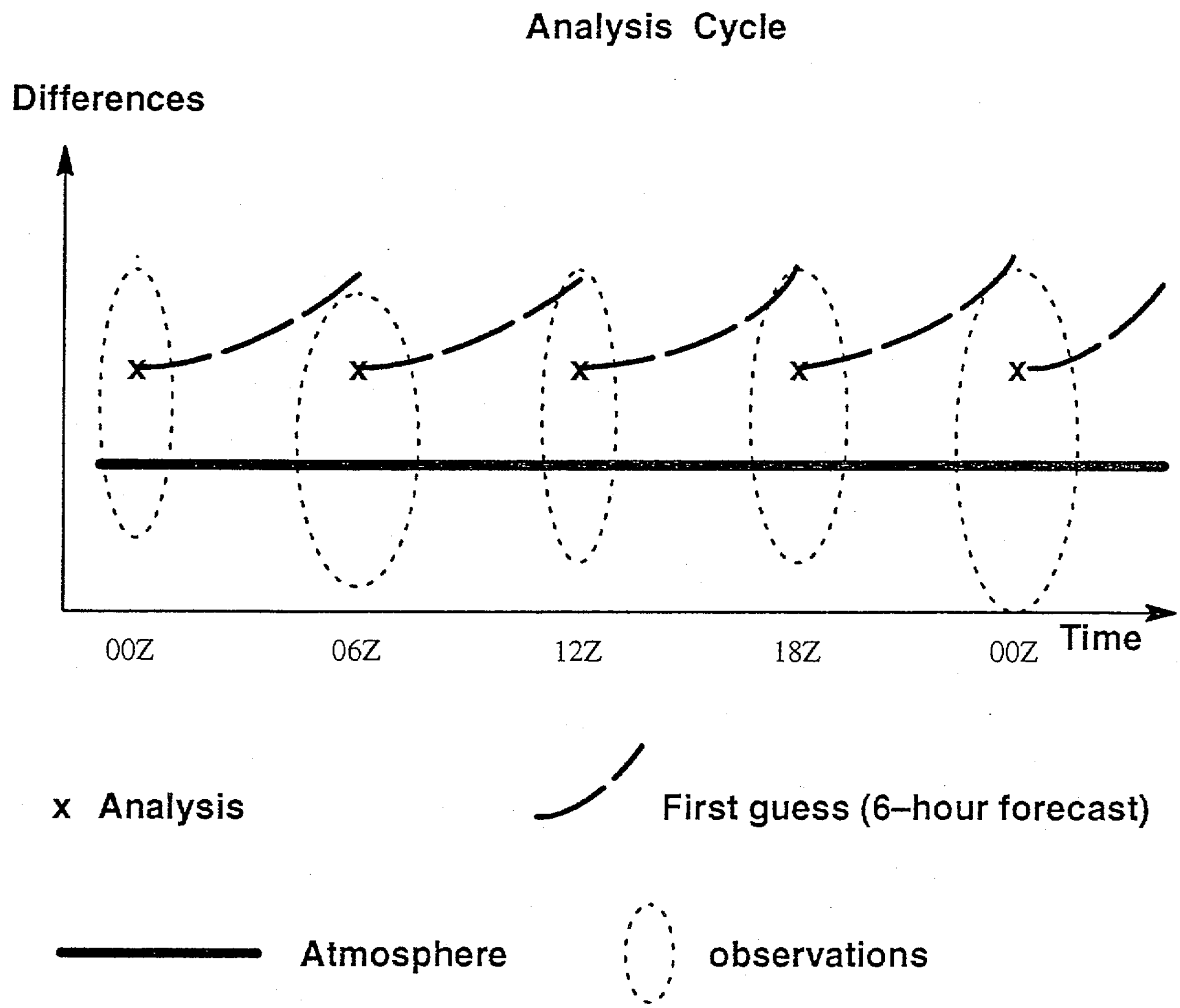

Analysis Cycle
Differences
00Z
06Z
12Z
18Z
00Z
Time
x Analysis
First guess (6–hour forecast)
Atmosphere
observations

Fig. 2

BREEDING OF GROWING MODES

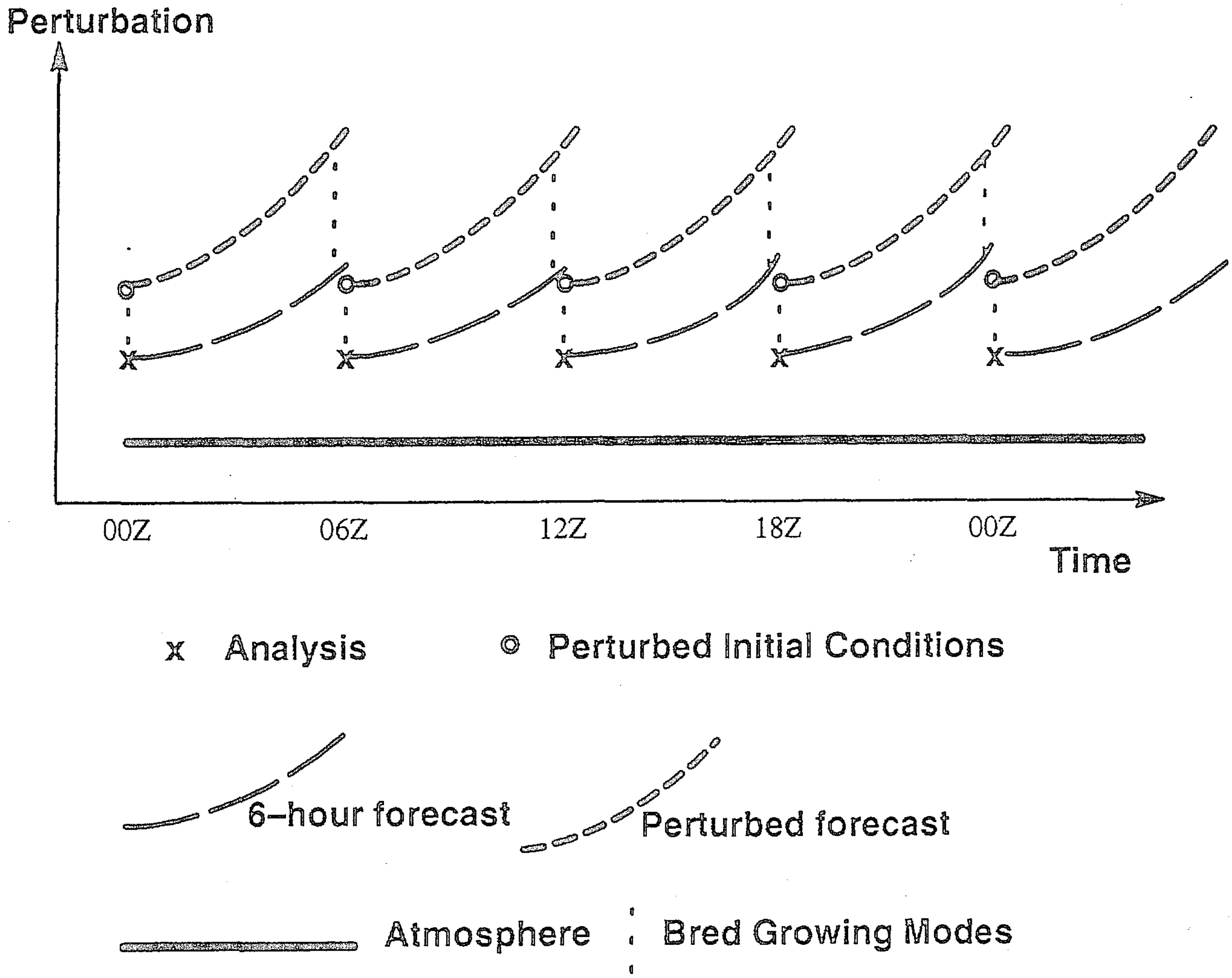


Fig. 3

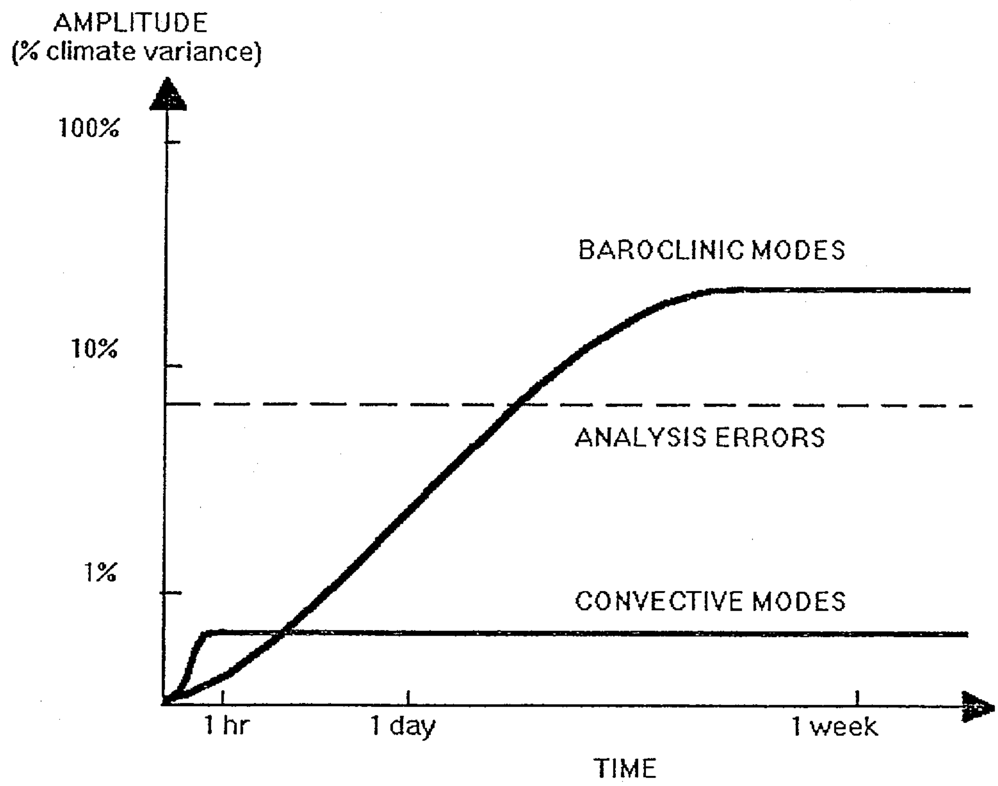

AMPLITUDE
(% climate variance)
100%
10%
1%
BAROCLINIC MODES
ANALYSIS ERRORS
CONVECTIVE MODES
1 hr
1 day
1 week
TIME

Fig. 4

SELF-BREEDING OF TWIN FORECASTS

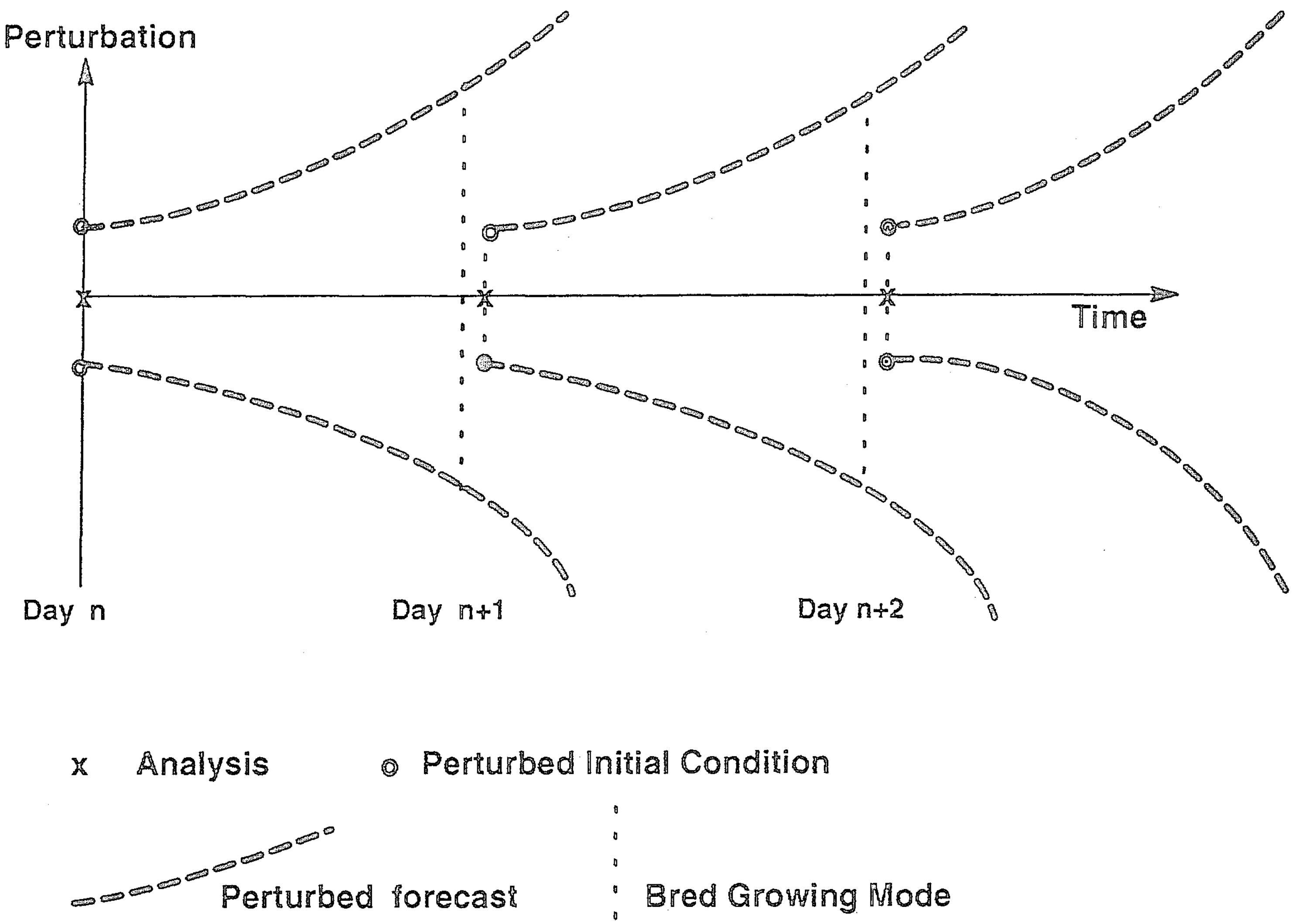


Fig. 5a

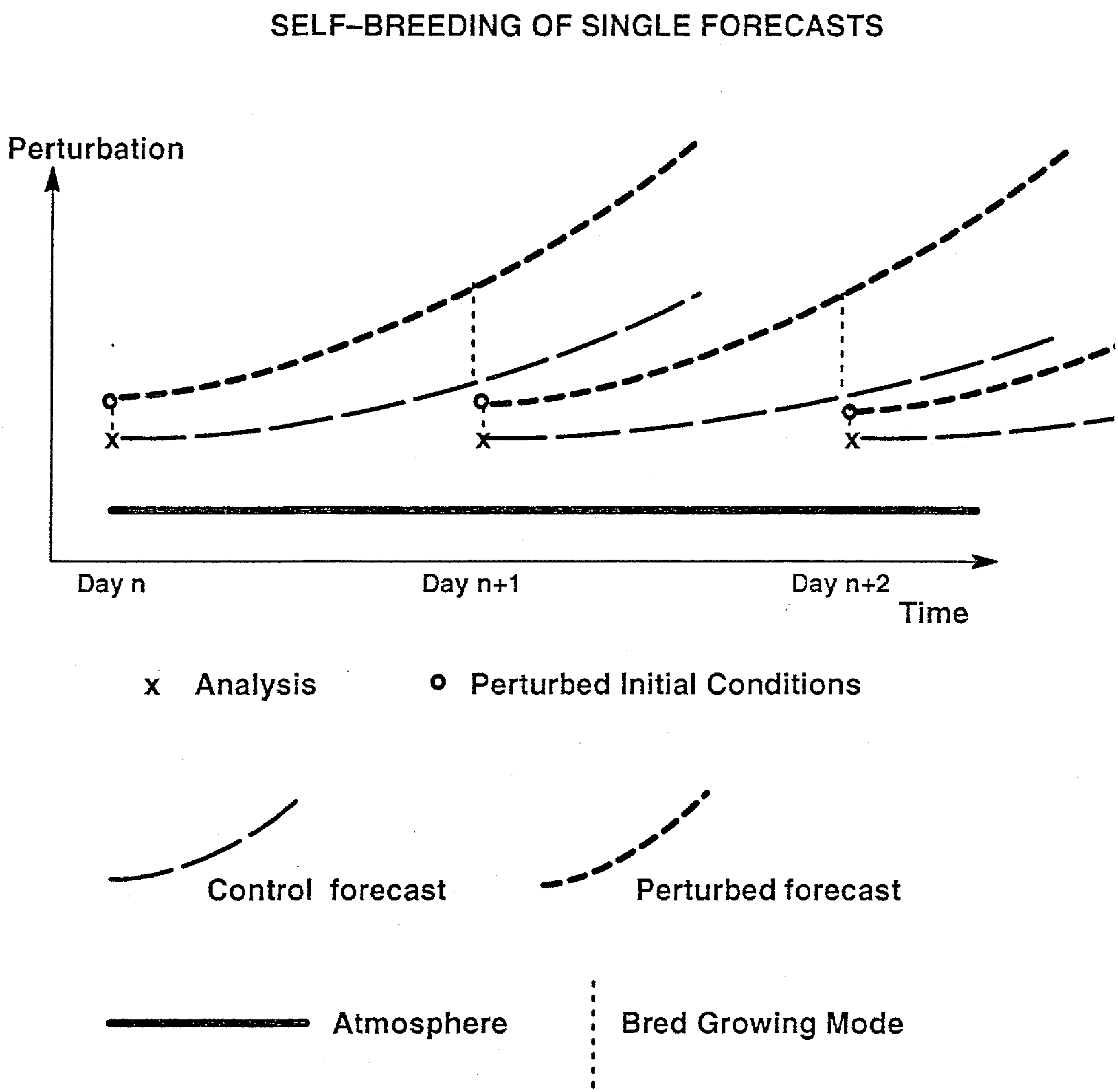
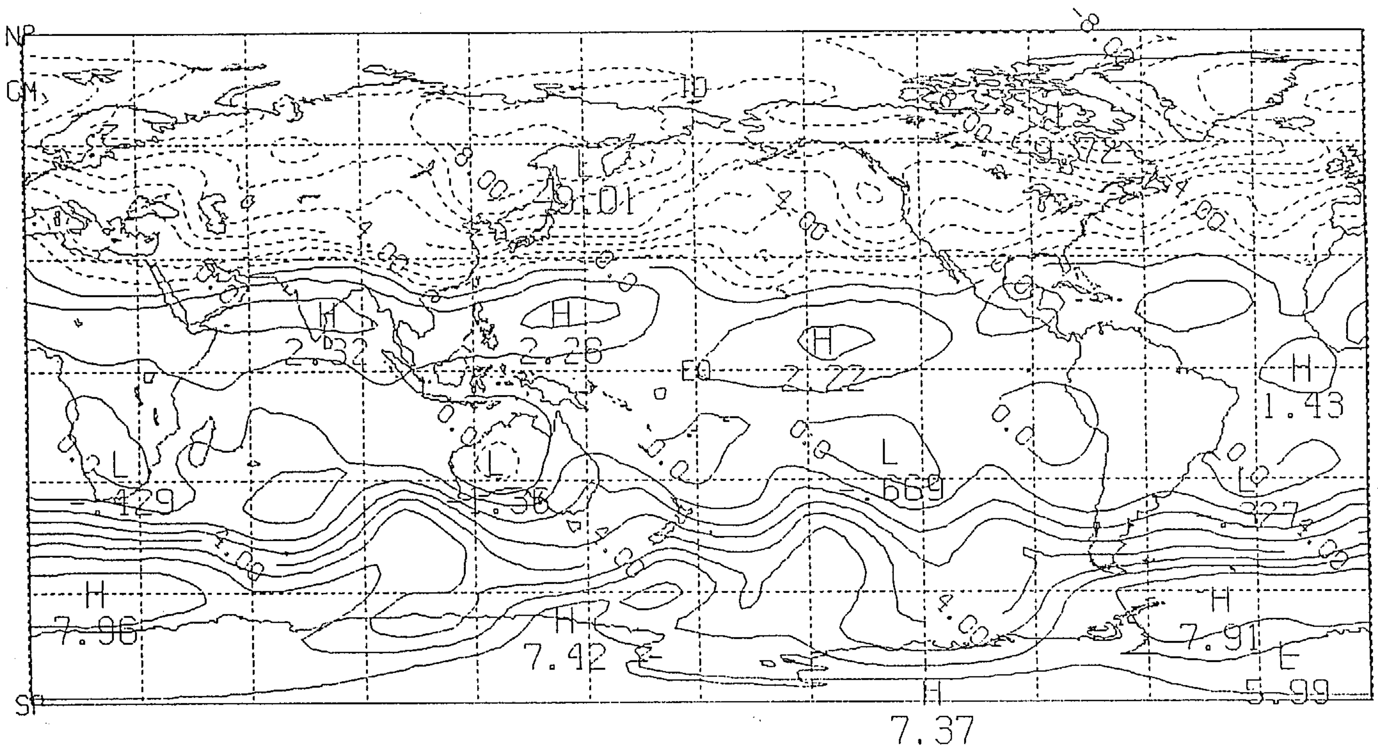

SELF–BREEDING OF SINGLE FORECASTS
Perturbation
Day n
Day n+1
Day n+2
Time
x Analysis
Perturbed Initial Conditions
Control forecast
Perturbed forecast
Atmosphere
Bred Growing Mode

Fig 5b



NP
GM
SP
EQ
H
2.32
H
2.26
H
2.22
H
1.43
L
9.01
L
-.429
L
-1.36
L
-.669
H
7.96
H
7.42
H
7.91
H
7.37
5.99

Fig. 6a

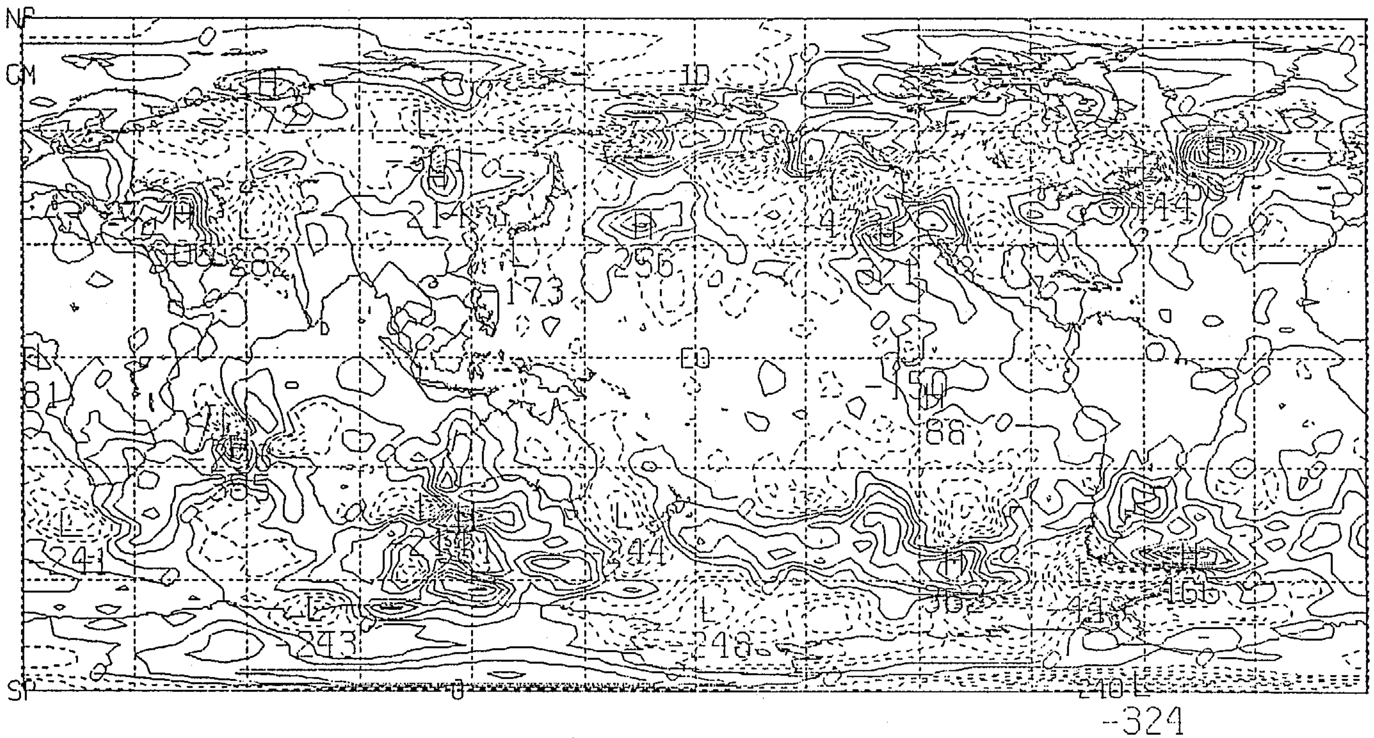
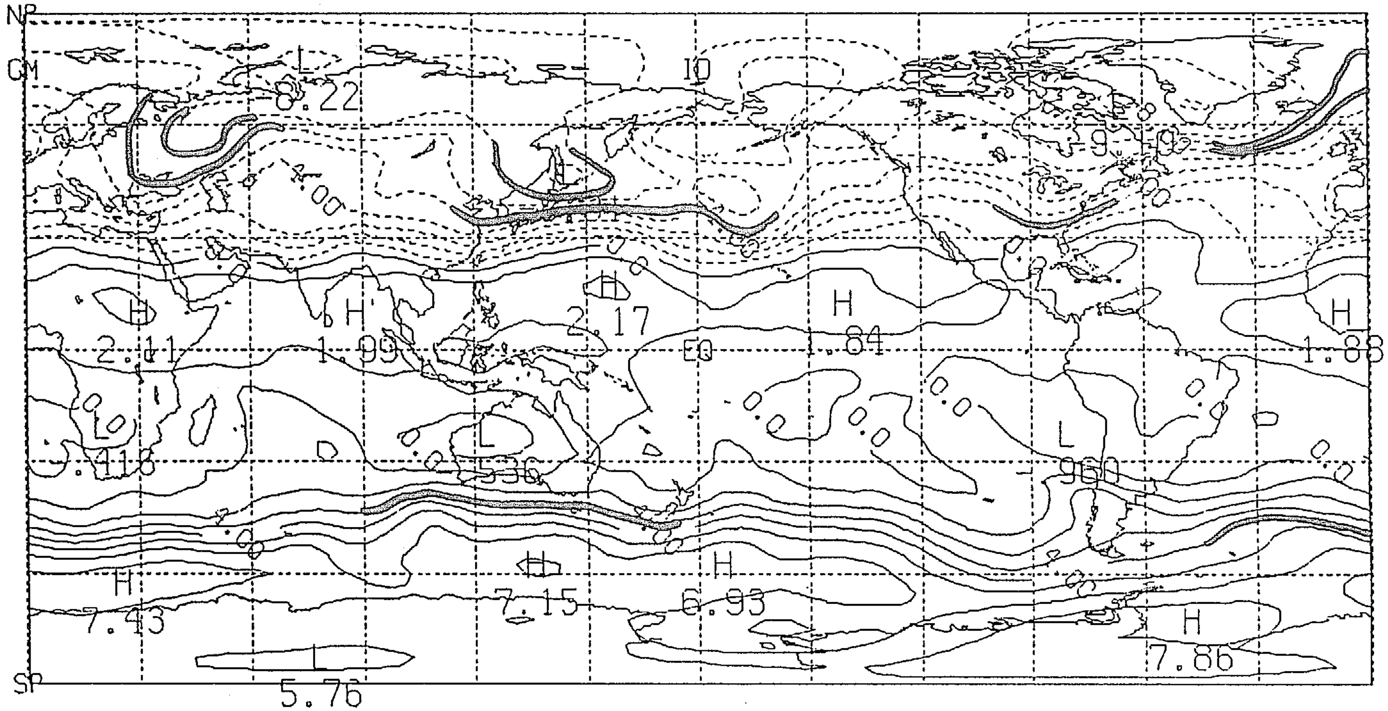


Fig. 6b



NP
GM
L
8.22
4.00
IO
H
2.17
EQ
H
2.11
H
1.99
H
1.84
H
1.88
0.0
L
.418
L
.536
L
960
H
7.43
H
7.15
H
6.93
H
7.86
L
5.76
SP

Fig. 7a

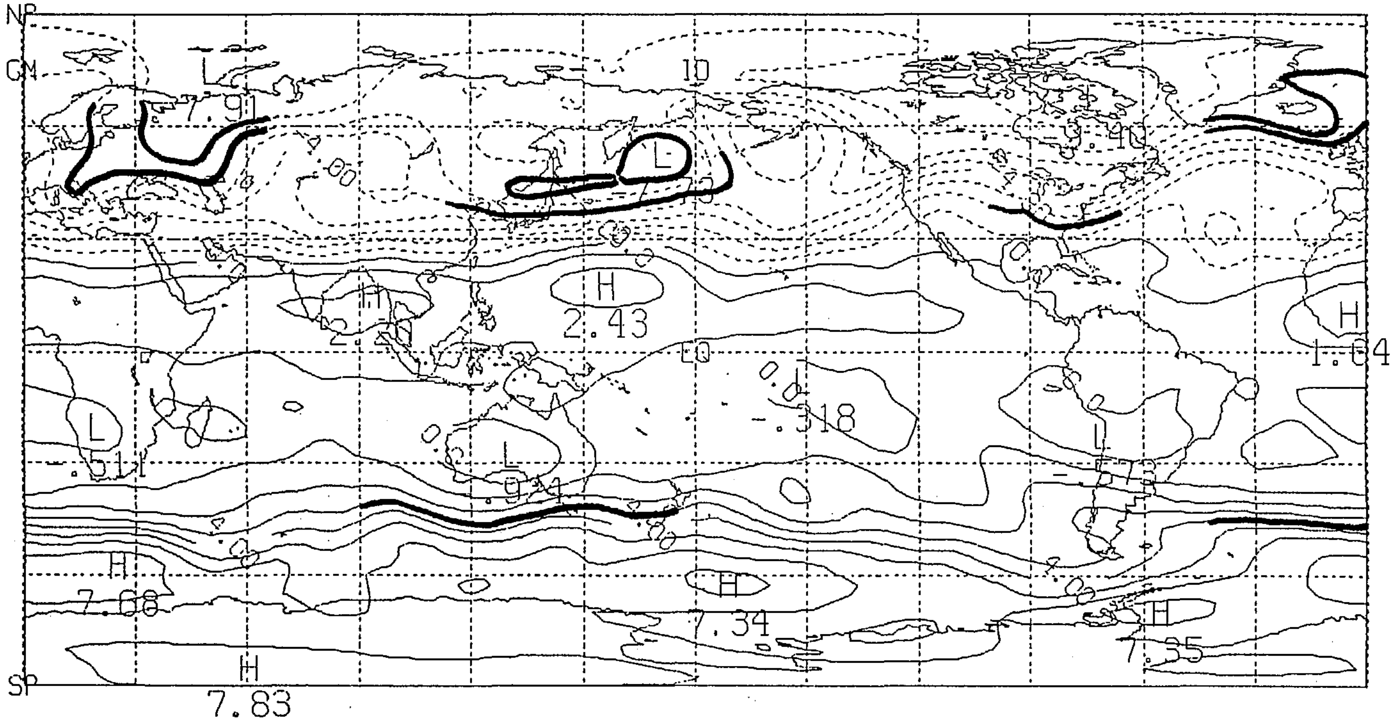


Fig. 7b

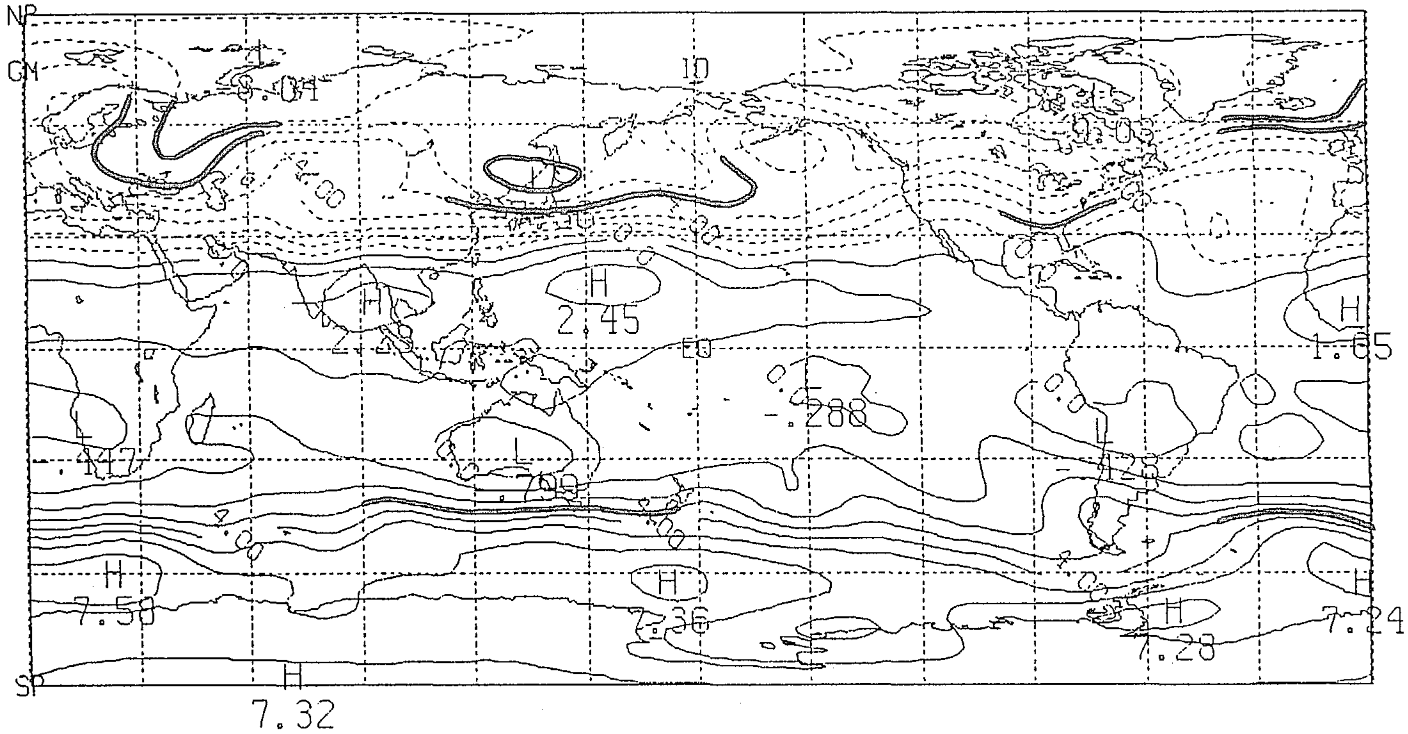


Fig. 7c

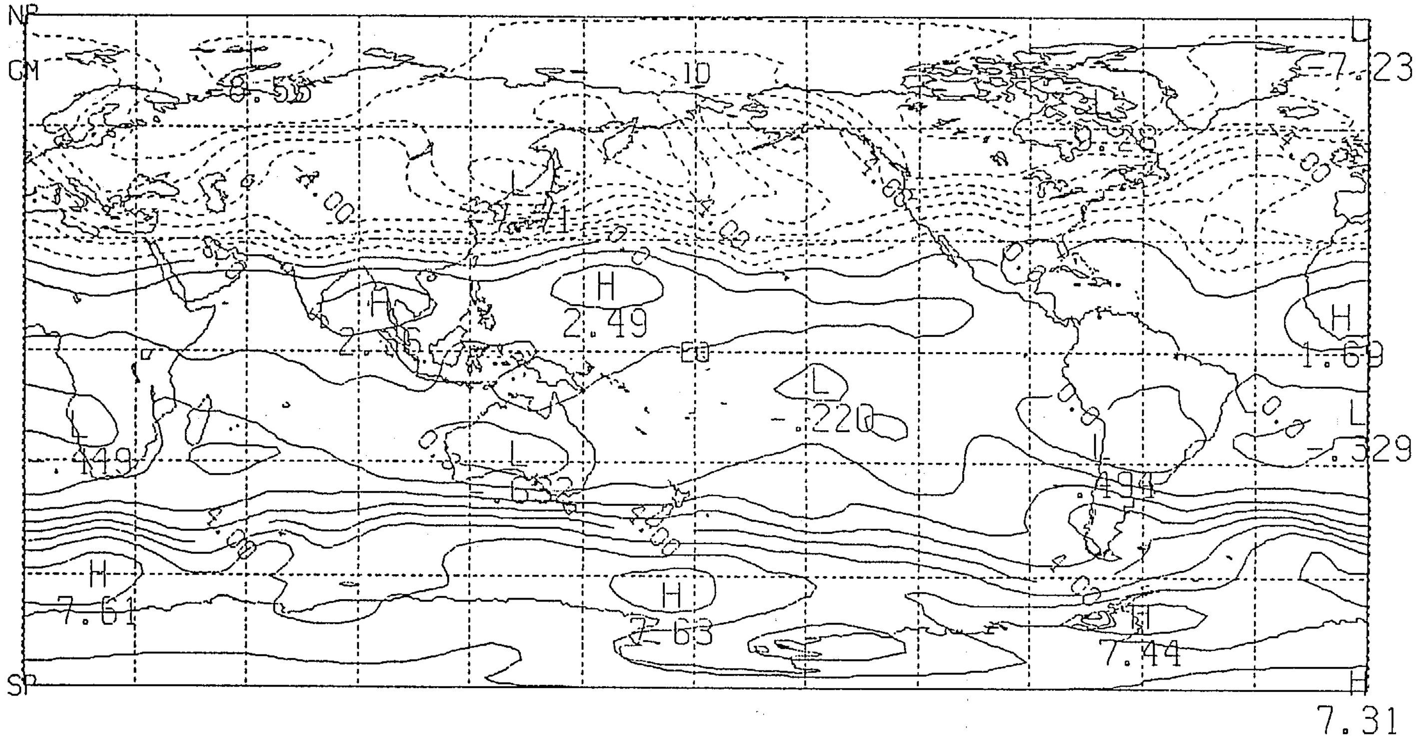


Fig. 8a

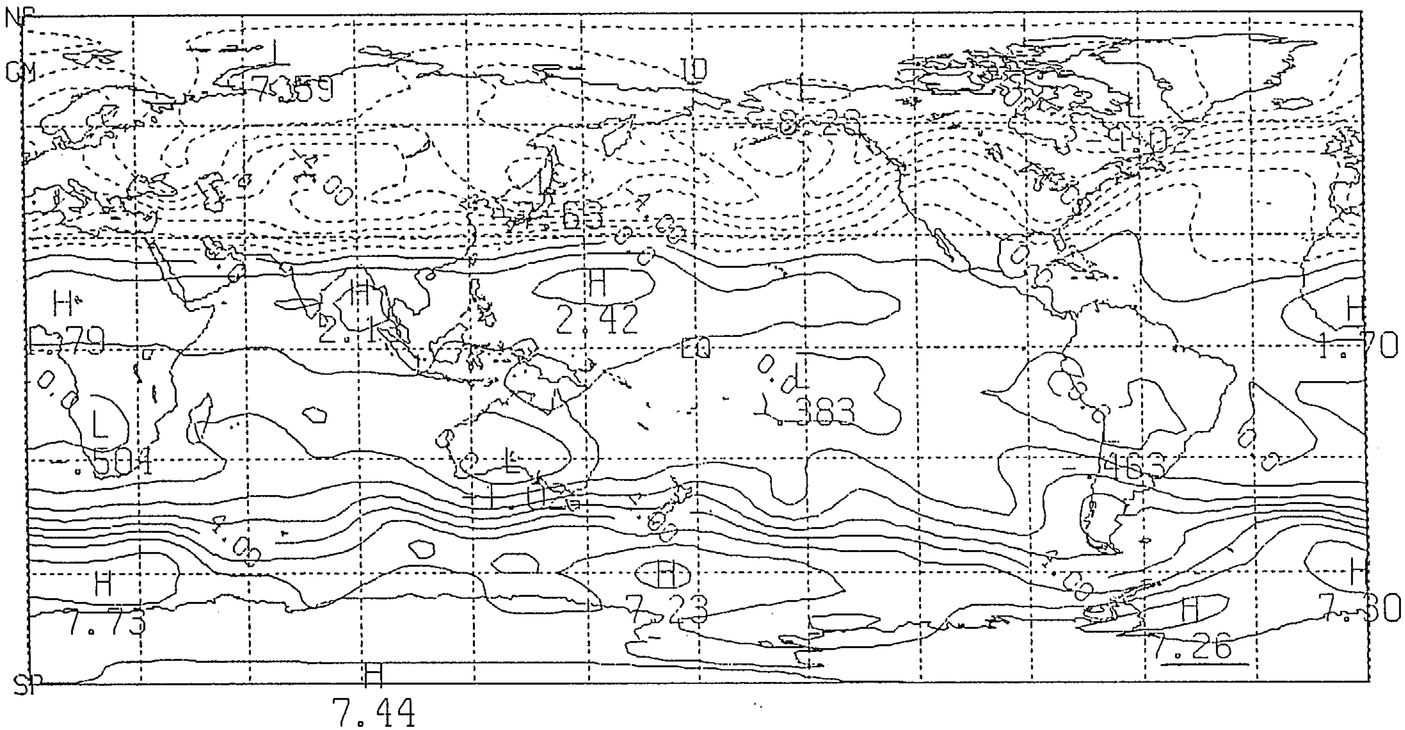

NP
GM
SP
EQ
L
7.59
L
8.28
L
9.02
7.63
H
2.42
H
2.1
H
1.79
H
1.70
L
.501
L
.383
.463
L
H
7.73
H
7.23
H
7.26
H
7.30
H
7.44

Fig. 8b